Political Activism and the Provision of Dynamic Incentives: Growing the Pie in the Battle for Redistribution

Antoine Camous, Russell Cooper*

January 18, 2021

Abstract

This paper studies the determination of income taxes in a dynamic setting with human capital accumulation. The goal is to understand the factors that support an outcome without complete redistribution, given a majority of relatively poor agents and the inability to commit to future taxes. All agents agree *ex ante* that limiting tax and transfers is beneficial but a majority favors large redistribution, *ex post*, at the time of the vote. In a political influence game, group activism limits the support for expropriatory taxation and preserves incentives. In some cases, the outcome corresponds to the optimal allocation under commitment.

Keywords: *activism, electoral competition, commitment, redistribution, human capital.*

JEL classification: D72, D74, E62, H31.

1 Introduction

Income distributions are asymmetric, with a majority of agents earning below average income. In contrast to predictions of simple models of electoral competition, labor income taxes are progressive but far from confiscatory. What factors constrain the relatively poor from expropriating the income and wealth of the relatively rich?

One approach to answering this question builds upon the *promise of upward mobility* (POUM): the poor of today recognize that they may be the wealthy of tomorrow, and accordingly refrain from voting for high redistribution rates.[1] In this setting, POUM works to alter the political preferences of the relatively poor by promoting the idea of a high degree of mobility within the income distribution.

Alternatives consider directly properties of the political process that prevent the demographic majority of the poor to translate into a political majority. *De jure* or *de facto* elements break the ideal of *one person, one vote* and

*Camous, Department of Economics, University of Mannheim, L7 3-5, 68131 Mannheim, Germany, camous@uni-mannheim.de. Cooper, Department of Economics, European University Institute, russellcoop@gmail.com. Thanks to Sebastian Findeisen, Andrea Mattozzi, Guido Tabellini, Fritz Sager, and seminar participants at University of Mannheim, European University Institute and Harvard CES. Comments from referees and our editor are gratefully acknowledged.

[1]The literature review discusses in detail these contributions along with papers that are directly linked to the main themes of our model.

provide higher relative influence to richer agents. Various institutions involved in the decision process contribute to eliminate highly redistributive outcomes by, in effect, reducing the political power of the relatively poor.

We propose an alternative mechanism that emphasizes the role of *group activism*.[2] Our notion of activism relates to the analysis of "political competition among pressure groups" conceptualized by Becker (1983): *political influence is not simply fixed by the political process but can be expanded by expenditures of time and money on campaign contributions, political advertising and in other ways that exert political pressure.* Our analysis has elements of the aforementioned approaches in that activism works through altering political preferences and reflects differential influence across income groups beyond demographics.

Our paper formalizes the idea of group activism and studies it in a multi-stage game played by political parties, individual citizens and coalitions that represent them. We find conditions such that the equilibrium outcome of the game entails less than full redistribution despite a majority of relatively poor voters. Because of a fiscal externality, all agents, from rich to poor, agree *ex ante* that limiting redistribution is beneficial but lack the ability to commit to future taxes. Consequently, the outcome of the vote might yield suboptimal high rates of redistribution. In contrast, group activism allows to shape voters' preferences and induces them not to support extreme tax proposals. Candidates, in deciding on their policy platforms, internalize how activism will tilt voters' preferences toward (or away from) their proposed redistribution rate. Accordingly, activism can eliminate the temptation of candidates to run on a platform of full redistribution.

We present these ideas in a model with an explicit human capital decision. Central to this environment is that incentives matter: individual agents make a long term investment choice that is forward looking and thus dependent on future tax rates. In this context, redistributive policies influence not just the allocation of the economic pie but also its size.[3]

There are two key ingredients in our model. First, the timing of tax policy relative to economic choices: the education decision is made prior to the choice of redistributive policies. Consequently, expectations about future taxes influence private investment, creating a familiar time consistency problem. Indeed, private agents understand that once human capital choices are made, tax policy will solely reflect redistributive gains. Standard electoral competition yields full redistribution of labor income, which is suboptimal. Thus the absence of commitment to tax policy makes its even more challenging to explain the limits to redistributive policies.

Second, there is activism itself. Activism embodies the idea that individuals, acting through groups, take joint actions to influence voters' political opinions towards their favorite tax platform.[4] This may take a variety of forms, ranging from direct persuasion within social circles to influencing public opinion through media and online campaigns. As the analysis makes clear, there are conflicts across influence groups: rich income activists support the low tax candidate while poor income agents promote the high tax one. The net effect of this influence game shifts permanently voters' perceptions of tax platforms. This mechanism is consistent with the idea that preferences are persistent and culturally ingrained into a political environment, as discussed in the literature section.

Again, timing is important: group activism contributes to curb redistribution if it is effective to shape preferences

[2]In our environment, neither the conditions associated to the POUM hypothesis hold - as formally studied by Bénabou and Ok (2001), nor there are other imperfections in the democratic process. These abstraction allows us to focus on the contribution of activism.

[3]As suggested by a referee, this interaction of timing and incentives could have equally been displayed in a model with endogenous labor supply. Complementary discussions in Appendix C rely on a model that, *inter alia*, includes a labor supply decision.

[4]This group decision component is also found in models of ethical voting in which voting participation is determined by groups of agents. See the literature review and Appendix C for further discussions of the differences between these approaches.

prior to the education decision. Indeed, all influence groups internalize the benefits of preserving dynamic incentives, hence the tax base for future redistribution. This provides a basis for relatively rich household groups to exert influence more in favor of low taxes and, at the same time, motivates the relatively poor to advocate less in favor of high taxes.[5] Hence, activism interacts directly with the choices of private agents through its effects on the probability distribution of future taxes. If, for example, activism reduces the probability of high taxes, agents retain an incentive to invest in human capital.

In equilibrium, activism eliminates the incentives of candidates to form a coalition around full redistribution. When the impact of activism is sufficiently strong on voters' preferences, the disciplining effect on candidates is powerful enough that the *ex ante* social optimum allocation is supported in equilibrium. Interestingly, there is no activism along the equilibrium path. The credible prospect of activism is enough to discipline candidates.[6]

Section 2 discusses the related literature. Section 3 presents the economic environment and derives social optimum levels of redistribution as policy benchmarks. Section 4 lays out the political economy game and studies how activism shapes economic outcomes under electoral competition. All derivations and proofs are detailed in an Appendix.

2 Literature

This section contains a review of the related literature from two perspectives. The first focuses on the political economy of redistribution. The second discusses analyses connected to the three central elements of our model: activism, group choice and timing. Papers that relate to model details and specific findings are discussed in the analysis.

2.1 Political economy of redistribution.

There is a large literature relating political institutions and redistribution, see Persson and Tabellini (2002) and Alesina and Giuliano (2011) for reviews. In the classic analysis by Meltzer and Richard (1981), democratic enfranchisement leads to higher tax rates, higher redistribution and lower inequality. Alesina and Ferrara (2005) presents empirical evidences that income level explains individual preferences for redistribution. The puzzle remains: why don't we observe higher levels of redistribution? The literature has considered two broad explanations that either limit the power of the relatively poor through various institutions or alter their preferences about redistribution.

First, democracy is imperfect: economic outcomes do not always reflect the demographic composition of the voting population. Examples include *one dollar, one vote* as in Karabarbounis (2011), or differential voting participation as in Bénabou (2000). Also, the political decision process involves various institutions which contribute to eliminate highly redistributive outcomes. Mattozzi and Snowberg (2018) study for instance the implications of legislative bargaining for redistribution. Developing an idea proposed by Roemer (1998), Fernández and Levy (2008) add another dimension to the policy decision: the vote is about the tax rate and the disposition of revenues. Though relatively poor agents may agree on the gains to taxing the rich, they may disagree about how to share

[5] Importantly, the poor need to participate in the activism game in favor of high taxes to balance the low tax preferences of the rich.
[6] In Section 4.4, we relax the assumption of vote share maximizing politicians and show that citizen candidates interested in the policy outcome interact in a way that generates activism along the equilibrium path.

the resulting proceeds. In this case, the rich may succeed in "dividing" the poor into smaller political groups and thus diffuse their collective power.

Alternatively, political preferences toward redistribution could be more elaborate than static economic considerations would suggest. Bénabou and Ok (2001) consider for instance an environment where the *promise of upward mobility* (POUM) influences the political preferences of the relatively poor against redistribution.[7] Two strict conditions need to be satisfied to support the POUM hypothesis under electoral competition. First, a change in the tax structure must be permanent, or at least difficult to undo. Second, the relatively poor median agent today votes against redistribution tomorrow if she believes she will become richer than average tomorrow.[8]

Relatedly, following Piketty (1995), some authors emphasize that voters are confronted with imperfect information on the origin of income inequality. This generates multiple systems of beliefs consistent with different levels of redistribution, as for instance in Alesina and Angeletos (2005), Bénabou and Tirole (2006) or Bénabou (2008).[9]

By construction, our analysis does not include any of these elements to precisely focus on the possible role of an alternative mechanism, namely activism.

2.2 Key model elements: activism, timing and group choice.

This section discusses key elements of our model in the context of the literature.

Activism. This paper relates to the large literature in social science on political activism. Building on the pioneering work by Olson (1965) on collective action and the paradox of coordination, Maisel, Berry, Edwards, and Schlozman (2010) argue that *interest groups are only a part of the set of organizations that represent collective interests in politics*: influence activities are undertaken by a multitude of organizations with diverse memberships, structures and purposes, generating a *heavenly chorus* of interests.

Our analysis retains some key characteristics of activism documented empirically: influence is effective and long lasting on voters' preferences.[10] For instance, Beck, Dalton, Greene, and Huckfeldt (2002) investigate whether limitations on acquiring and processing information could expose individual preferences *to being swamped by readily available information sources.* They document in the context of the 1992 US Presidential election how exposure to *partisan* information influences voting behaviors. Their findings support the idea that voters' policy preferences are shaped by social context and continuous information rather than specific electoral communication or factual news exposure.

This is further supported by Pennec and Pons (2019). Using a panel dataset of policy preferences and vote choice during electoral campaigns in OECD countries, they find that changes in vote choice during electoral seasons are mostly driven by changes in beliefs about candidates and policy salience, while policy preferences remain stable

[7] Alesina, Stantcheva, and Teso (2018) document how beliefs about intergenerational mobility affect preferences for redistribution using cross-country surveys.

[8] For this condition to be plausible, the income distribution should display negative skewness in the future, which is at odds with empirical income distributions across countries. In the absence of these conditions, then the relatively poor will favor the immediate and complete taxation of the rich.

[9] Note that asymmetric information only makes fully redistributive tax schedules costly to implement, because incentives must be provided for agents to report their income truthfully. Bierbrauer and Boyer (2013) and Bierbrauer and Boyer (2015) study the efficiency of political competition under this informational friction.

[10] For a review of the response of politicians to persuasive communication between the media and voters, see DellaVigna and Gentzkow (2010).

over time. The social origin and permanence of policy preferences is highlighted by Alesina and Fuchs-Schündeln (2007): they contrast East and West German policy preferences to investigate whether *45 years of heavy state intervention and indoctrination instill in people the view that the state is essential for individual well-being.* Their findings confirm that preferences are long lasting and that individual political experience shapes economic beliefs - for instance toward redistribution.

Timing. Our approach emphasizes the interaction between dynamic incentives and the lack of commitment to future policies. A well-known example arises in Fischer (1980) where investment in capital goods precedes the determination of capital taxation creating a familiar commitment problem. Several studies consider the implications of a lack of commitment on the political economy of dynamic taxation.[11] These include for instance Acemoglu, Golosov, and Tsyvinski (2010), where elections operate as a disciplining device on politicians to implement tax schedules similar to the one that would prevail under commitment.

Group Decisions. Our analysis introduces activism as a group decision interacting with policy choice protocols. Becker (1983) presents a related theory, where competition among pressure groups to curry favor from politicians leads to beneficial economic outcomes. Political environments with strategic group choice also include the ethical voter's models by Coate and Conlin (2004) or Feddersen and Sandroni (2006), where voters' turnout is endogenous: partisan groups weigh the costs of taking part to the vote against the benefits to their affiliated group.[12] In these environments, the expression of policy preferences reflects the extensive margin of participation to voting, while our mechanism, activism, emphasizes an intensive margin channel, where interest groups compete for influencing voters.[13]

Bierbrauer, Tsyvinski, and Werquin (2017) combine elements of the ethical voter model to study taxation. The participation margin induces, for instance, left leaning parties not to propose high tax rates, to demobilize rich agents from voting. In contrast, our focus underlines the dynamic benefits of group choice in an environment with a lack of commitment. An essential difference concerns the timing of policy choice: in our model taxes are set after the education choice while in Bierbrauer, Tsyvinski, and Werquin (2017) taxes are set prior to labor supply. This difference matters particularly for understanding the channels through which group choice impacts private decisions.[14] Indeed, activism operates on the probabilistic outcome of the vote through its long lasting effect on the preferences of all voters, not just those that belong to the group engaging in activism. In the ethical voter model, in contrast, no such spillovers exist as the voter participation decisions are naturally group specific. Together, these elements connect all voters to a common goal: preserving a tax base from the risk of redistributive coalitions, so that in equilibrium there is output to redistribute. Our analysis provides conditions such that the political game with group choice supports the efficient level of redistribution.

[11] Farhi, Sleet, Werning, and Yeltekin (2012) or Scheuer and Wolitzky (2016) study nonlinear capital taxation under exogenous and partial commitment and various political protocols. In the absence of commitment, full redistribution is the equilibrium outcome in their environments.

[12] Levine and Mattozzi (2020) develop this framework to study social norm enforcement of voter turnout with costly peer punishment.

[13] An alternative expression of relative intensity of policy preferences could be votes trading or shifting as in Casella (2005). See Casella and Macé (2020) for a review.

[14] This is brought out clearly in the model and discussions of Appendix C.

3 Environment

Consider a two period $t = 1, 2$ economy populated by a continuum of agents. Agents at $t = 1$ differ in ability $\theta \sim \log\mathcal{N}(m, \sigma^2)$, with cumulative and probability distribution functions noted respectively $F(\theta)$ and $f(\theta)$. They make a long run investment choice, e.g. in education, which influences next period's income. At $t = 2$, agents are subject to productive idiosyncratic shocks. Taxes and transfers are applied to period 2 gross income, and agents consume net income. The economic channels of fiscal policy are multiple: it *redistributes* income across the population, provides *insurance* against idiosyncratic shocks and *distorts dynamic choices*. The specific functional forms adopted here are chosen to highlight transparently the interplay between political institutions, growth and redistribution.

3.1 Individual dynamic choice

An agent with ability θ at $t = 1$ invests in education e to maximize lifetime utility:

$$\max_{e} \; \log(c) + \beta E_{z,\tau}\big(\log(c')\big), \tag{1}$$

subject to:

$$c = \theta - e \qquad \text{budget constraint } t = 1 \tag{2}$$

$$\theta' = z\theta^{\alpha}e^{\delta} \qquad \text{gross income } t = 2 \tag{3}$$

$$c' = \theta'^{1-\tau}\bar{\theta}'^{\tau} \qquad \text{net income } t = 2. \tag{4}$$

An agent with ability θ is referred to as a type θ agent.

Equation (2) is period 1 budget constraint: gross income is either consumed or spent on education. The wage is normalized at unity so that labor income is equated to initial ability θ. Equation (3) is the dynamic evolution of human capital and hence income. Both current ability θ and education choice e determine future income θ' up to an idiosyncratic shock $z \sim \log\mathcal{N}\big(-\frac{w^2}{2}, w^2\big)$. The parameters $\alpha > 0$ and $\delta \geq 0$ measure respectively depreciation of human capital and return to education.[15]

Fiscal policy in period 2 is captured by an isoelastic tax function, see (4), where $\bar{\theta}'$ characterizes the tax base for fiscal interventions as discussed below. This form is common in the literature on progressive labor taxes: the higher the redistribution rate $\tau \in [0, 1]$, the lower the dispersion of net income.[16] Redistribution also provides insurance against individual variation in productivity z.

The expectation operator in (1) reflects uncertainty over the magnitude of taxes and transfers τ. The institutional structure that determines the period 2 tax rate will be a key element in the analysis.

[15] By design, individual education is not directly influenced by the human capital decisions of others. This allows to isolate the interaction of agents through fiscal policy.

[16] Bénabou (2000), Bénabou (2002) and Heathcote, Storesletten, and Violante (2017), among others, adopt a similar tax function. Under this specification, agents report their true income to the tax authority. *Ex post* incentives to misreport income are not included in this analysis, allowing us to focus on the *ex ante* adverse incentive effects of taxation on human capital accumulation.

The optimal education choice satisfies:

$$e(\theta, \bar{\tau}) = \epsilon(\bar{\tau})\theta, \tag{5}$$

where $\bar{\tau}$ is the expected average redistribution rate, and $\epsilon(\bar{\tau}) \equiv \frac{\beta\delta(1-\bar{\tau})}{1+\beta\delta(1-\bar{\tau})}$ is an education rate common to all agents. In the limit case of extreme redistribution, i.e., $\bar{\tau} = 1$, there is no private return to education, so that in period 2 agents have zero income and thus zero consumption. This highlights the incentive effects of the labor tax.

Evolution of the income distribution. Gross income at $t = 2$ is log-normally distributed, with mean m' and standard deviation σ' given by:

$$m' = (\alpha + \delta)m + \delta \log\big(\epsilon(\bar{\tau})\big) - \frac{w^2}{2}, \tag{6}$$

$$\sigma'^2 = (\alpha + \delta)^2\sigma^2 + w^2. \tag{7}$$

Fiscal intervention. Through taxes and transfers determined by τ, the fiscal intervention is purely redistributive. The critical income level $\bar{\theta}'$ sorts agents in net contributors $\theta' > \bar{\theta}'$ and net beneficiaries $\theta' \leq \bar{\theta}'$. It satisfies:

$$\log(\bar{\theta}') = m' + \frac{\sigma'^2}{2}(2 - \tau). \tag{8}$$

Note the multiple influences of redistribution on $\bar{\theta}'$. First, average log-income m' is a function of education and the expected tax rate $\bar{\tau}$, as explicit in (6). This expression makes clear that the model contains a fiscal externality: if all agents in the economy increase their education level, then $\bar{\theta}'$ will increase, and if $\tau > 0$ so will the consumption of all agents.[17] Also, $\bar{\theta}'$ is directly decreasing in τ: the share of the population that pays more taxes than it receives transfers is increasing in the magnitude of the redistributive program.

3.2 Individual bliss policies

To highlight the influence of redistribution on individual preferences and welfare, we contrast individual bliss policies at two points in time: before agents form an education choice at $t = 1$, and after, at $t = 2$, when idiosyncratic uncertainty remains.[18] These policies produce normative benchmarks against which the outcomes of political protocols are evaluated.

At $t = 1$, before agents choose education. Let $V_1(\theta, \tau)$ be the value function of an agent of type θ evaluating a rate of fiscal redistribution τ. Combining the optimal education choice (1) with lifetime utility (5), it reads:

$$V_1(\theta, \tau) = \log\big(\theta - \epsilon(\tau)\theta\big) + \beta\Big((1-\tau)\big(\underbrace{\alpha\log(\theta) + \delta\log\big(\epsilon(\tau)\theta\big) - \frac{w^2}{2}}_{=E_z\left(\log(\theta')|\theta\right)}\big) + \tau\log(\bar{\theta}')\Big). \tag{9}$$

[17] Though no single agent takes this into account in making its education choice, both *ex ante* bliss tax policies and group choices will internalize this effect,

[18] This timing allows the exogenous component of mobility to shape individual preferences. It is maintained throughout the analysis.

There are two ways in which τ influences $V_1(\theta,\tau)$. First, the individual education decision $e(\theta,\tau)=\epsilon(\tau)\theta$ is sensitive to τ, as explicit in (5). Second, the fiscal tax base, captured by the break-even income level $\bar{\theta}'$ (8), responds to the tax rate as well.

The favorite redistribution rate $\tau^*(\theta)$ of a type θ agent is the tax rate $\tau \in [0,1]$ that solves $\frac{dV_1(\cdot)}{d\tau}=0$. It is implicitly given by:

$$\beta(\alpha+\delta)\big(m-\log(\theta)\big)+\beta\big(\underbrace{(\alpha+\delta)^2\sigma^2+w^2}_{=\sigma'^2}\big)(1-\tau)+\beta\delta\frac{\tau}{\epsilon(\tau)}\frac{d\epsilon(\tau)}{d\tau}=0. \tag{10}$$

The first two terms in this expression capture the relative support for *redistribution*. At least all agents with income below (log) median level m benefits (on average) from redistributive policies. Also, an increase in individual risk w^2 generates higher desire for *insurance* against idiosyncratic shocks z. The third term is the elasticity of the education rate $\epsilon(\tau)$ to redistribution τ. It is negative and captures the joint gain from preserving individual *dynamic incentives* and thus the tax base against distortionary redistribution. It is straightforward to show that $\tau^*(\theta)<1$ is unique and decreasing in θ: higher income agents prefer lower rates of redistribution.[19]

At $t=2$, after individual choices. Let $V_2(\theta,\tau|\epsilon)$ be the value function of a type θ agent after $t=1$ consumption and education choices, before a realization of idiosyncratic risk z.[20]

$$V_2(\theta,\tau|\epsilon)=\beta\Big((1-\tau)\big(\underbrace{\alpha\log(\theta)+\delta\log(\epsilon\theta)-\frac{w^2}{2}}_{=E_z\left(\log(\theta')|\theta\right)}\big)+\tau\log(\bar{\theta}')\Big). \tag{11}$$

At that stage, education levels are no longer sensitive to tax policies τ: the education rate ϵ is a sufficient statistic to describe individual education $e=\epsilon\theta$ and the aggregate tax base.[21] The favorite redistribution rate $\tau^d(\theta)$ is either an interior solution to $\frac{dV_2(\cdot)}{d\tau}=0$:

$$\beta(\alpha+\delta)\big(m-\log(\theta)\big)+\beta\big(\underbrace{(\alpha+\delta)^2\sigma^2+w^2}_{=\sigma'^2}\big)(1-\tau)=0, \tag{12}$$

or the corner solution $\tau^d(\theta)=1$ for lower income agents.[22] As education is realized, there is no term capturing *dynamic* implications of redistribution, in contrast to (10). This expression reflects otherwise the relative preferences for redistribution, decreasing with income, and insurance against idiosyncratic shocks.

Figure 1 represents bliss policies before and after education, respectively $\tau^*(\theta)$ and $\tau^d(\theta)$. Favorite redistribution rates are decreasing in income θ. When agents internalize the effect of taxes on incentives, they favor lower rates than they would after education is made: $\tau^*(\theta)\le\tau^d(\theta)$ for all θ. Finally, after education, a majority supports

[19] For formal derivations, see Appendix A. This expression reflects in part our assumptions on preferences and distributions. But the generic trade-offs embedded in the interactions between dynamic choices and redistributive policies generalize to less restrictive preference and technology specifications.

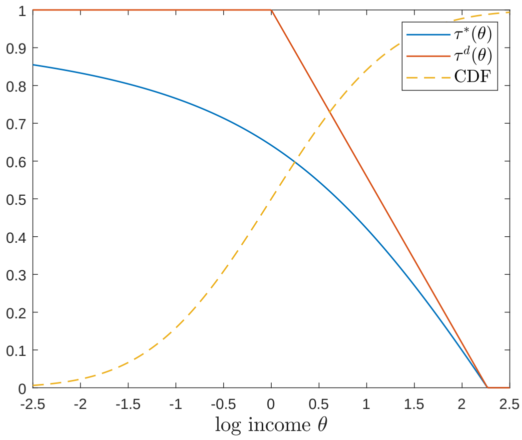
[20] Given that the ranking of income before the realization of idiosyncratic uncertainty is maintained over time, we continue to order agents by initial income θ and cumulative distribution function $F(\theta)$.

[21] Also, the break even income level $\bar{\theta}'$ is sensitive to τ insofar that it sorts agents between contributors and beneficiaries of the tax program, given average log income m'.

[22] Note that bliss policies $\tau^d(\theta)$ are only a function of income, not education ϵ: the size of the tax base is irrelevant and only redistributive conflicts determine bliss policies.

complete redistribution.

Figure 1: Individual Bliss Policies



This figure represents bliss policies before and after education, as a function of (log) income. Favorite rates are decreasing in income and agents support higher levels of redistribution once education is made. The yellow dashed line shows the underlying distribution of log-income. [Illustrative calibration: $\beta = 0.96$, $\alpha = \delta = 0.3$, all other parameters set to 1]

3.3 Normative benchmarks

These bliss policies facilitate the derivation of two normative policy benchmarks, distinguished by the timing of the policy choice relative to individual education decisions. Assume first that the tax policy is chosen by a benevolent planner with commitment at $t = 1$. That is, taxes are chosen prior to the education choice and *ex post* are not subject to change. Formally, a benevolent policy maker chooses a tax rate τ^* to maximize expected utility over the population:

$$\max_{\tau} \int_{\theta} V_1(\theta, \tau) dF(\theta). \tag{13}$$

In contrast taxes τ^d set without commitment are chosen in period 2, after the education decision ϵ.

$$\max_{\tau} \int_{\theta} V_2(\theta, \tau|\epsilon) dF(\theta). \tag{14}$$

The following proposition characterizes these policy benchmarks.

Proposition 1. *A benevolent planner:*

(i) with commitment sets $\tau^* < 1$, where τ^* is the solution to

$$\beta\big(\underbrace{(\alpha+\delta)^2\sigma^2 + w^2}_{=\sigma'^2}\big)(1-\tau) + \beta\delta\frac{\tau}{\epsilon(\tau)}\frac{d\epsilon(\tau)}{d\tau} = 0, \tag{15}$$

(ii) without commitment implements a fully redistributive tax rate $\tau^d = 1$.

The terms in (15) highlight how $\tau^* < 1$ balances (average) preferences for *redistribution / insurance* and *incentives*. The first term captures redistribution across ability θ and insurance against human capital risk z: the optimal rate of redistribution is increasing in income inequality σ^2 and idiosyncratic risk w^2.[23] The second term captures the negative effect of higher taxes on human capital accumulation, i.e., on the tax base. The magnitude of this effect is parameterized by the return to education δ.[24]

In contrast, if taxes are set once agents have made their education choice, then *dynamic incentives* are no longer internalized. Policy choice is only driven by redistributive considerations, hence the policy maker implements full redistribution. In this second case, private agents have no incentives to invest in education, and the economy collapses.

4 Political activism and redistributive policies

We now consider a political decentralization of policy choice, where competing candidates propose tax rates to maximize vote share. Our political environment combines activism and probabilistic voting. Probabilistic voting follows Lindbeck and Weibull (1987): agents evaluate candidates based upon their preferred tax rates and another dimension that reflects political preferences. Political activism has an impact on the voting outcome through this second dimension.

In the game, education decisions are taken prior to policy choice. This timing reflects the wedge between individual long term education decisions and repeated elections, that give rise to the risk of excessive redistribution. Actually, in the absence of group activism, full redistribution is the equilibrium outcome under electoral competition.

We then introduce activism. Agents organize in large and self-interested groups and decide on joint actions to influence voters' political preferences. Importantly, activism is effective to influence electoral competition if it takes place prior to the education choice: individual agents under the influence of activist groups adjust their beliefs about future taxes and undertake education accordingly. Hence, political activism operates on the outcome of the vote by forging voters' preferences and preserving incentives to invest in human capital.

Importantly, candidates appreciate that their choice of policy platforms can elicit a response through the level and direction of political activism. As the influence of activism grows, it restores efficiency by providing incentives to candidates not to run on highly redistributive platforms.[25]

4.1 Timing of the game

The timing highlights both that human capital is determined prior to the vote on taxes and how activism influences the election outcome. Formally, the sequence of events is:

i. Choice of platforms: two office seeking candidates from competing parties L and H propose redistributive platforms $\tau_l \leq \tau_h$.

[23] If redistribution over initial income θ were allowed, this would decrease initial income inequality and the optimal rate of redistribution. But the commitment tension at the heart of the mechanism would be maintained.

[24] The overall effect of δ on τ^* is ambiguous, since it contributes both to higher return to education but also to income dispersion.

[25] We provide in Appendix C a complement to the analysis: a labor supply choice substitutes education in a simplified version of our game to further isolate the importance of timing and the channels of influence of activism.

ii. Group activism: income groups engage in a game of influence.

iii. Individual choice at $t = 1$: agents, given their type θ, choose consumption and education e.

iv. Political preferences: individuals are subject to idiosyncratic and aggregate political preference shocks.

v. Vote: given policy platforms and political preferences, agents participate in a majority election and the winning candidate takes office.

vi. Realizations of individual income shock z, tax and transfer and $t = 2$ consumption.

This timing calls for some comments. First, individual income uncertainty realizes after the vote, to give a chance for *insurance* and *upward mobility* to influence the vote against the most redistributive policy platforms. Second, the vote takes place after the education choice, precisely to investigate whether highly redistributive platforms would arise in a game where the tax base is predetermined to the election. Finally, we introduce activism prior to the individual choice, allowing it to shape jointly individual decisions and political preferences. As politicians are forward looking, this allows activism to influence candidates' policy platforms.

4.2 Electoral competition without activism

In order to isolate the effects of activism from the other elements of the political protocol, the first step is to characterize the political equilibrium of this game without activism. Absent group activism, policy platforms (τ_l, τ_h) are decided by office seeking candidates anticipating individual education decisions and the outcome of the probabilistic vote. The following exposition derives from the sequential nature of the game.

Voting outcome. Given policy platforms (τ_l, τ_h) and education rate ϵ, individuals trade off political and economic preferences and cast their vote sincerely. They evaluate policy platforms $\tau_l \leq \tau_h$ according to the value function $V_2(\theta, \tau|\epsilon)$ along with the realizations of idiosyncratic χ and aggregate ψ political preference shocks for candidate L. A type θ agent with education $e = \epsilon\theta$ votes for party H if and only if:

$$V_2(\theta, \tau_h|\epsilon) > V_2(\theta, \tau_l|\epsilon) + \chi + \psi. \tag{16}$$

As in Persson and Tabellini (2002), these shocks are distributed as:

$$\chi \sim U\Big(-\frac{1}{2\phi}, \frac{1}{2\phi}\Big) \qquad\qquad \psi \sim U\Big(-\frac{1}{2\Psi}, \frac{1}{2\Psi}\Big). \tag{17}$$

They differ only because the average of the idiosyncratic shock χ across the population is zero, while ψ is common across agents.[26]

Given a realization of aggregate preference ψ, let $\chi(\theta, \psi)$ be the *swing voter* for agents with income θ: type θ agents vote for party H if and only if $\chi \leq \chi(\theta, \psi)$. From (16),

$$\chi(\theta, \psi) = V_2(\theta, \tau_h|\epsilon) - V_2(\theta, \tau_l|\epsilon) - \psi = \Delta V_2(\theta) - \psi, \tag{18}$$

[26] Though these shocks determine political preferences for candidate L, they impact the voting outcome symmetrically. As shown in Appendix B, the results are not sensitive to the mean zero assumption, or to the strict majority requirement introduced below.

where $\Delta V_2(\theta)$ is the economic gain (or loss) to agents with initial income θ of τ_h over τ_l:

$$\Delta V_2(\theta) = \beta(\tau_h - \tau_l)\Big[(\alpha+\delta)\big(m - \log(\theta)\big) + \frac{\sigma'^2}{2}(2 - \tau_h - \tau_l)\Big]. \tag{19}$$

This expression makes clear that $\chi(\theta, \psi)$ and $\Delta V_2(\theta)$ do not depend on the actual education rate ϵ, as individual preferences are only driven by distributional conflicts at this stage. The vote share for party H within group θ and across the population are:

$$\pi_{\theta,h}(\psi) = \int_{-\frac{1}{2\phi}}^{\chi(\theta,\psi)} \phi dj = \phi\Big(\chi(\theta,\psi) + \frac{1}{2\phi}\Big) \qquad\qquad \pi_h(\psi) = \int_\theta \pi_{\theta,h}(\psi) dF(\theta). \tag{20}$$

In a majority system, the probability p_h that the candidate from party H wins the election is:

$$p_h = P\big(\pi_h(\psi) \geq 1/2\big). \tag{21}$$

Combining previous expressions:[27]

$$p_h = \frac{1}{2} + \Psi\beta(\tau_h - \tau_l)\frac{\sigma'^2}{2}(2 - \tau_h - \tau_l) \geq \frac{1}{2}. \tag{22}$$

This expression highlights the tendency of the population to lean toward the most redistributive platform τ_h. When elections take place after the education choice, then a pure redistributive conflict drives economic preferences of agents. As illustrated in Figure 1, the positive skewness of the income distribution provides a majority mass of the population benefiting from high redistribution rates.

Choice of platforms. A candidate from party H seeking to maximize its probability of winning chooses to campaign on a redistributive program τ_h that solves:

$$\max_{\tau_h \in [0,1]} p_h(\tau_l, \tau_h), \tag{23}$$

where $p_h(\tau_l, \tau_h)$ is given by (22). The first order condition is:

$$\frac{dp_h(\cdot)}{d\tau_h} = \Psi\beta\sigma'^2(1 - \tau_h) = 0. \tag{24}$$

$\tau_h = 1$ is a dominant strategy. The candidate from party L maximizes $p_l(\tau_l, \tau_h) = 1 - p_h(\tau_l, \tau_h)$, and again $\tau_l = 1$ is a dominant strategy.

Proposition 2. *The outcome of the game without group activism is full redistribution* $\tau^p = 1$.

Intuitively, the election takes place after education choices have been formed. Hence despite redistributive conflicts, the median income agents with average political preferences supports unconditionally a platform of full

[27] Naturally $0 \leq p_h \leq 1$. We abstract from assumptions on parameters and restrictions on choices that are not relevant to characterize equilibrium outcomes.

redistribution. In this case, dynamic incentives cannot be preserved and the economy collapses[28] This no longer holds once we introduce activism as an avenue to influence political preferences and electoral outcomes.

4.3 Electoral Competition with Activism

We now consider the equilibrium of the game when agents can form influence groups and actively support candidates and platforms.[29] By activism, we mean non-pecuniary contributions aimed to influence political preferences of voters.

We analyze activism intensity chosen by income groups and its impact on the voting outcome and equilibrium policy platforms. The analysis takes as given group membership and focuses on the intensive margin of group specific contributions.[30] Groups are large and thus internalize the effects of their actions on the voting outcome.[31]

Activism impact on the voting outcome. After the announcement of policy platforms (τ_l, τ_h), each income group θ decides on (non-pecuniary) activism intensity $A_\theta^h \geq 0$ and $A_\theta^l \geq 0$ to promote candidates and their economic platforms. Group activism influences political preferences of **all** voters, not just those belonging to the group initiating activism.

At the time of the vote, these choices are all given. The vote reflects the economic valuations of policy platforms along with political preferences. Note $A^i = \int_\theta A_\theta^i d\theta$ the aggregated influence of activism for each candidate $i \in \{l, h\}$. Given $\left(A^l, A^h\right)$, aggregate preference shock ψ, education rate ϵ and policy platforms (τ_l, τ_h), an agent with initial income θ and preference shock χ votes for party H if and only if:

$$V_2(\theta, \tau_h|\epsilon) > V_2(\theta, \tau_l|\epsilon) + \chi + \psi + \gamma(A^l - A^h), \tag{25}$$

where $\gamma \geq 0$ measures how activism influences political preferences. The probability p_h that party H wins the election and τ_h is implemented reads:

$$p_h = \frac{1}{2} + \Psi \int_\theta \Delta V_2(\theta) dF(\theta) + \Psi\gamma(A^h - A^l). \tag{26}$$

Superficially, the term $\gamma(A^l - A^h)$ seems only to shift the distribution of the aggregate political preference shock. This misses two critical dimensions of the analysis: the choice of platforms is made anticipating the levels of activism and individual education decisions that reflect the expected outcome of the vote. In this way, the prospect of political influence impacts both the tax base for fiscal interventions and the electoral platforms of candidates.

[28] Probabilistic voting per se does not curb the tendency of the population to support high rate of redistributions once the tax base is fixed. To be clear, if the vote were to take place <u>before</u> education, then electoral competition would yield the efficient outcome τ^* characterized in Proposition 1.This is formally discussed in Appendix B.2.

[29] The equilibrium concept is now two stage Nash equilibrium: first candidates choice of platforms and then activism effort.

[30] For expositional clarity, we interpret the continuous probability distribution function $f(\theta)$ as a probability mass function that reflects the relative size of income group θ, and hence their *market power* in the activism game. In Appendix B.4, we detail an adjustment of the environment where we consider a partition of the population in $N \geq 2$ groups that engage in an otherwise exactly similar activism game. Note that the political economy literature has adopted similar hypotheses in various environments, for instance when considering a continuum of voters where participants vote as if they were pivotal. This underlies for instance the notion of "conditional sincerity" by Alesina and Rosenthal (1996), further adopted by Caselli and Morelli (2004) and Mattozzi and Merlo (2008). We thank Andrea Mattozzi for helping us to clarify these subtleties at the very beginning and for the most recent revision of this project.

[31] This is a feature that is shared by other group-based decision processes, such as the ethical voter model, but the channels of influence differ. See Appendix C for an explicit comparison of the two approaches.

Activism choice. Every income group takes part in the influence game, with the goal of promoting its own economic interests. Activism takes the form of effort, whose cost appears as a utility loss.

Given competing platforms $\boldsymbol{\tau} = (\tau_l, \tau_h)$ and all other group activism decisions $\{A^l_{-\theta}, A^h_{-\theta}\}$, income group θ decides on total group contributions $\left(A^l_\theta, A^h_\theta\right)$:

$$\max_{A^l_\theta, A^h_\theta \geq 0} f(\theta)V_1(\theta, \boldsymbol{\tau}) - \frac{1}{2}\Big((A^l_\theta)^2 + (A^h_\theta)^2\Big). \tag{27}$$

Here $f(\theta)$ is the mass of agents in the income group θ and $V_1(\theta, \boldsymbol{\tau})$ is the expected value of a type θ household prior to the education choice, given competing political platforms.[32] It is similar to (9) but captures uncertainty over the tax rate $\tau \in (\tau_l, \tau_h)$:

$$V_1(\theta, \boldsymbol{\tau}) = E_\tau \left\{ \log\big(\theta - \epsilon(\bar{\tau})\theta\big) + \beta\big((1-\tau)\alpha\log(\theta) + \delta\log(\epsilon(\bar{\tau})\theta) - \frac{w^2}{2} + \tau\log(\bar{\theta}')\big) \right\}. \tag{28}$$

Activism internalizes the effect of contributions on the education rate $\epsilon(\bar{\tau})$ and the outcome of the vote, captured by the probability p_h that party H defeats party L, see (26). The income break even level $\bar{\theta}'$, which reflects the tax base for fiscal intervention, is a function of both $\bar{\tau}$ and τ, as explicit in (8). Finally, the group cost in (27) is assumed to be quadratic. One interpretation is there is a cost of organization associated with activism that depends on the contributions of all members of the group.[33]

The first order condition for $A^i_\theta \geq 0$, $i \in (l, h)$, is

$$f(\theta)\frac{dV_1(\theta)}{dA^i_\theta} = A^i_\theta. \tag{29}$$

The sensitivity of group θ welfare to activism is:[34]

$$\frac{dV_1(\cdot)}{dA^i_\theta} = \pm\Psi\gamma\beta(\tau_h - \tau_l)\Big[(\alpha+\delta)\big(m - \log(\theta)\big) + \frac{\sigma_2^2}{2}(2 - \tau_h - \tau_l) + \delta\bar{\tau}\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})}\Big]. \tag{30}$$

The first two terms reflects the relative position of group θ in the income distribution and the preferences for redistribution, as in (10).

The last term captures the effects of group contributions on the probability of high taxes and, through the expected tax rate, on the collective accumulation of human capital $\epsilon(\bar{\tau})$, i.e., the tax base for fiscal interventions.[35] This term is central to understand the outcome of the activism game: regardless of an agent's position in the income distribution, this effect pushes in the direction of low taxes since all agents prefer a larger tax base.

This last term is also the locus of strategic interactions, i.e., conflicts across the groups. Again, the decision to support a low or high tax candidate as well as the magnitude of group contributions depend on the probability

[32] To be clear, A^i_θ is the activism effort collectively exerted by group θ, hence the value function $V_1(\theta, \boldsymbol{\tau})$ is weighted by group size $f(\theta)$. Also, the program (27) omits uncertainty regarding future political taste shocks. This does not change the analysis beyond overburdening notations.

[33] If the individual cost to agent j with income θ of contributing $a^{i,j}_\theta$ is $\frac{a^{i,j}_\theta}{2}a^i_\theta f(\theta)$, where a^i_θ is the average group contribution, then total group costs is $\frac{{A^i_\theta}^2}{2} = \frac{(a^i_\theta f(\theta))^2}{2}$: individual cost is linear, but increases with group size and average contribution. Appendix C recasts our main result under a general increasing and convex cost function.

[34] In this expression, $\pm = \mathbb{1}_{i=h} - \mathbb{1}_{i=l}$.

[35] Formally, $\bar{\tau} = p_h\tau_h + (1-p_h)\tau_l$.

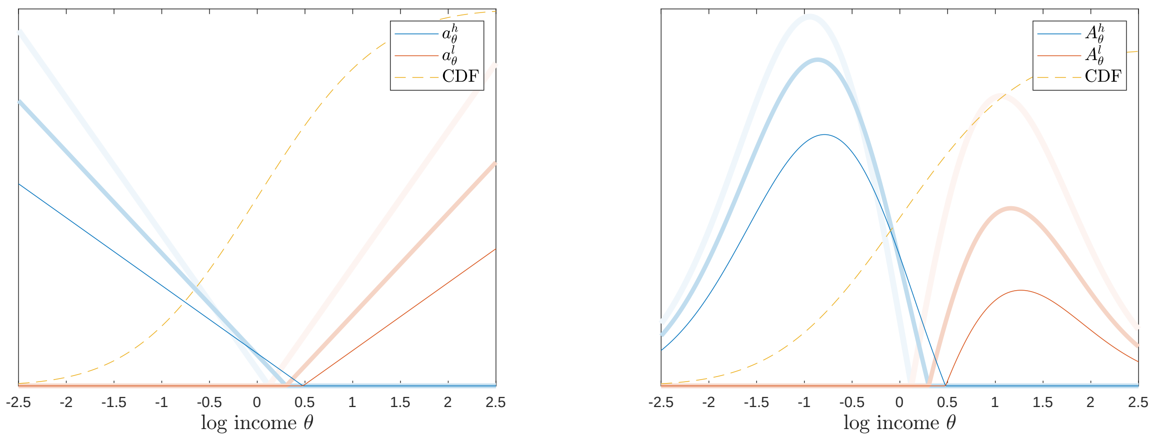
p_h, which hinges on overall activism across the population, as explicit in (26). As long as $\tau_l < \tau_h$, all groups are active, and promote only one candidate: low income agents support only the champion of high taxes, while high income agents only the candidate from party L. Importantly, the split of the population is endogenous. All groups with initial income $\theta < \hat{\theta}$ contribute exclusively for τ_h, where the cut-off income level $\hat{\theta}$ is given by:

$$\log(\hat{\theta}) = m + \frac{1}{\alpha + \delta}\Big[\frac{\sigma'^2}{2}(2 - \tau_h - \tau_l) + \delta\bar{\tau}\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})}\Big]. \tag{31}$$

Figure 2 reports activism per capita $a^i_\theta = A^i_\theta / f(\theta)$ and total group activism A^i_θ for some policy platforms $\tau_l < \tau_h$. Activism per capita is increasing in the income difference $|\hat{\theta} - \theta|$. Aggregate group effort is not monotonic, because of the relative size $f(\theta)$ of each income group. Finally, an increase in activism technology γ has two effects: it increases returns to activism but modifies the cut-off $\hat{\theta}$, i.e., it changes the composition of the population that promotes one candidate or the other. Aggregating group contributions (29), one gets:

$$A^h - A^l = \Psi\gamma\beta(\tau_h - \tau_l)\Big[\frac{\sigma'^2}{2}(2 - \tau_h - \tau_l) + \delta\bar{\tau}\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})}\Big]. \tag{32}$$

Figure 2: Activism Contribution



(a) Activism per capita (b) Total group activism

This figure represents activism per capita (left panel) and total group contributions (right panel) given two policy platforms $\tau_l = 1/4$ and $\tau_h = 3/4$. Lighter lines correspond to higher values of γ. The dashed yellow line represents the distribution of (log) income. [Illustrative calibration: $\beta = 0.96$, $\alpha = \delta = 0.3$, all other parameters set to 1]

Lemma 1. *Given competing tax platforms (τ_l, τ_h), there is a unique Nash equilibrium of the activism subgame. The induced probability that the high tax candidate wins the election is given by:*

$$p_h(\tau_l, \tau_h) = \frac{1}{2} + \Psi\beta(\tau_h - \tau_l)\Big[(1 + \Psi\gamma^2)\frac{\sigma'^2}{2}(2 - \tau_h - \tau_l) + \Psi\gamma^2\delta\bar{\tau}\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})}\Big]. \tag{33}$$

The probability of high taxes takes into account the aggregation of group activism (32) and reflects the resolution of conflicts over redistribution. Importantly, this expression highlights the interactions of activism and human capital accumulation in reducing the probability of high tax rates. In the absence of a return to education, i.e. $\delta = 0$, $p_h \geq \frac{1}{2}$ as in (22). Only when $\delta > 0$ and $\gamma > 0$ does the last term in (33) reduce the probability of high taxes.

Choice of platforms. How does activism influence political competition and equilibrium incentives? At the initial stage of the game, each candidate decides on its economic policy platform anticipating the effects of activism on voters' political preferences and the outcome of the vote. Formally, candidate from party H sets τ_h given τ_l to maximize (33):

$$\max_{\tau_h \in [0,1]} p_h(\tau_l, \tau_h). \tag{34}$$

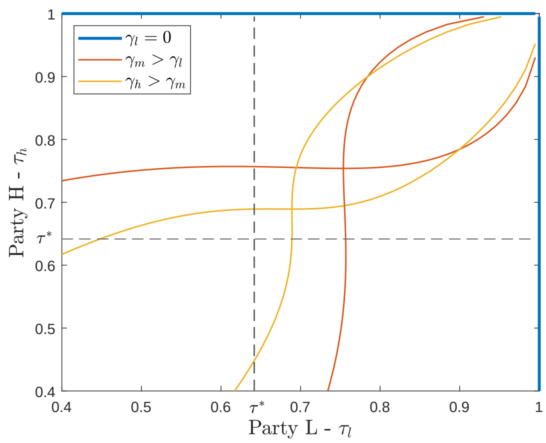
The first order condition leads to:

$$(1 + \Psi\gamma^2)\sigma'^2(1 - \tau_h) + \Psi\gamma^2\delta\bar{\tau}\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})} + (\tau_h - \tau_l)\Psi\gamma^2\delta\frac{d\mathcal{E}(\bar{\tau})}{d\bar{\tau}}p_h(\tau_l, \tau_h) = 0, \tag{35}$$

where $\mathcal{E}(\bar{\tau}) = \bar{\tau}\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})} < 0$ is the elasticity of the education rate to the expected redistribution rate.

The platform choice internalizes the effect of the tax rate on the outcome of the activism subgame. Again, it highlights the interplay between activism and collective dynamic incentives. In the absence of dynamic choice, i.e. when $\delta = 0$, activism and associated conflicts are irrelevant since $\tau_h = 1$ is a dominant strategy, as in (24). In contrast, when $\delta > 0$, then activism induces strategic interactions across candidates. Figure (3) represents the best response functions of each candidate for different level of activism technology γ.

Figure 3: Candidates Best Response



This figure represents the best response function of candidates at the platform choice stage, without activism ($\gamma_l = 0$) and with varying degrees of activism ($\gamma_h > \gamma_m > 0$). [Illustrative calibration: $\beta = 0.96$, $\alpha = \delta = 0.3$, all other parameters set to 1]

Lemma 2. *There is a unique and symmetric Nash equilibrium $\tau^p = \tau_l = \tau_h$ that satisfies:*

$$\Big(1 + \frac{1}{\Psi\gamma^2}\Big)\sigma'^2(1 - \tau) + \delta\tau\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})} = 0. \tag{36}$$

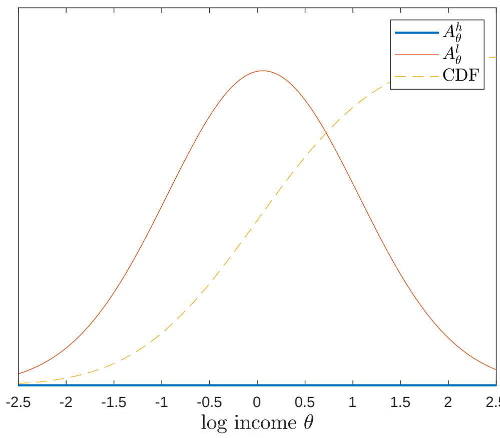
Activism does not imply that parties have an incentive to differentiate from each other in equilibrium, but as long as there are positive returns to education $\delta > 0$, the equilibrium no longer coincides with full redistribution.

Proposition 3. *The equilibrium rate of redistribution τ^p differs from $\tau = 1$ if and only if $\delta > 0$ and $\gamma > 0$. Further, it is decreasing in activism intensity γ, and in the limit case where $\gamma = +\infty$, the political game with activism implements the socially desirable level of redistribution τ^*.*

Along the equilibrium path of this game, there is no political activism since $\tau_l = \tau_h = \tau^p$. Still activist groups stand ready to influence voters' political preferences if a candidate were to deviate from the equilibrium platform. If, for instance, the candidate from party L deviates from τ^p and runs on $\tau_l < \tau^p$, then relatively poor agents would engage into activism to convince voters away from this platform. And vice versa. This is the disciplining effect of activism on candidates.

In particular, full redistribution cannot be the outcome of electoral competition under the influence of group activism. To see why, consider a proposed equilibrium in which both candidates offer the platform $\tau = 1$. Now imagine that candidate L deviates and proposes $\tau_l = 1 - \iota$. As illustrated in Figure 4, then *almost* all the population promotes the low tax rate candidate at the activism stage.[36] The platform deviation hence increases the probability that candidate L wins the election. Hence, $(\tau_l, \tau_h) = (1, 1)$ cannot be an equilibrium of the political game with activism.

Figure 4: No Full Redistribution under Activism



This figure represents total group activism for competing platforms $\tau_l = 0.95$ and $\tau_h = 1$: the population *almost* unanimously exerts influence to support the lower tax platform. Accordingly, full redistribution cannot be an outcome of the game with activism. [Illustrative calibration: $\beta = 0.96$, $\alpha = \delta = 0.3$, $\gamma > 0$ and all other parameters set to 1]

The intensity of the activism technology γ influences the equilibrium rate of redistribution τ^p via the conflictual behavior across groups: for higher γ, the disciplining effect of activism is stronger. Again, the response of education to taxes, parameterized by δ is necessary for this channel.

Overall, activism contributes to support socially desirable policies in dynamic environments with a lack of commitment. It matters because agents, regardless of their position in the income distribution, appreciate the

[36] As ι gets to zero, then the income cut-of $\hat{\theta}$ (31) that splits the population in supporters of each candidate tends to zero.

social benefit of human capital accumulation and thus a large tax base for fiscal interventions. So, as long as there is a response of education to activism's effect on expected taxes, the bias towards redistribution under majority voting is, at least partially, redressed.

4.4 Extensions

This section studies two key extensions. What is the effect of asymmetric activism technology on equilibrium outcomes? And, how do office-seeking candidates also interested to the policy outcome interact?

Asymmetric influence. The equilibrium outcome with activism characterized in Proposition 3 does not rely on exogenous asymmetry across groups to move away from full redistribution.[37]

In this section, we show that the introduction of asymmetry in activism technology might further reduce the likelihood of high taxes. Formally, we consider a situation where income groups differ in their technology to influence political preferences. The cost of contributing is decreasing in α_θ, with $\int_\theta \alpha_\theta d\theta = 1$. Given policy platforms $\boldsymbol{\tau} = (\tau_l, \tau_h)$ and other groups' contributions $\{A^l_{-\theta}, A^h_{-\theta}\}$, the activism choice of income group θ is the solution to:

$$\max_{A^l_\theta, A^h_\theta \geq 0} f(\theta) V_1(\theta, \boldsymbol{\tau}) - \frac{1}{2\alpha_\theta}\Big((A^l_\theta)^2 + (A^h_\theta)^2\Big). \tag{37}$$

Let $\rho = \text{cov}\big(\alpha_\theta, \log(\theta)\big)$ be the covariance between income level and influence technology. The first order condition for $A^i_\theta \geq 0$ then reads:

$$\alpha_\theta f(\theta) \frac{dV_1(\theta, \boldsymbol{\tau})}{dA^i_\theta} = A^i_\theta. \tag{38}$$

Group contribution is increasing in influence technology α_θ. As $E\big(\alpha_\theta \log(\theta)\big) = m_1 + \rho$, the probability $p_h(\tau_l, \tau_h)$ of party H win is then implicitly defined by:

$$p_h(\tau_l, \tau_h) = \frac{1}{2} + (\Psi\gamma)^2 \beta(\tau_h - \tau_l)\Big[(1 + \frac{1}{\Psi\gamma^2})\frac{\sigma'^2}{2}(2 - \tau_h - \tau_l) + \delta\bar{\tau}\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})} - (\alpha + \delta)\rho\Big]. \tag{39}$$

A Nash equilibrium of the game across office seeking candidates $\tau^p = \tau_l = \tau_h$ is then the solution to:

$$\Big(1 + \frac{1}{\Psi\gamma^2}\Big)\sigma'^2(1 - \tau) + \delta\tau\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})} - (\alpha + \delta)\rho = 0. \tag{40}$$

In that context, $\frac{d\tau^p}{d\rho} < 0$: the larger the influence of high income groups, the lower the equilibrium rate of redistribution. Accordingly, a positive covariance between income level and activism technology can either compensate a low aggregate activism technology γ and bring the equilibrium level of redistribution toward the optimal level or can tilt the policy rate towards partisan level of redistributions that are too low.

[37] In contrast, within a probabilistic voting model, Persson and Tabellini (2002) consider asymmetry in the mass of swing voters across the population as a necessary condition to avoid full redistribution: when high income agents are more attractive targets to candidates, this reduces the likelihood of high taxes. See Appendix B.1 for a formal exposition of this result within our model. The same results hold if we allow differential voting participation across income groups, as in Bénabou (2000).

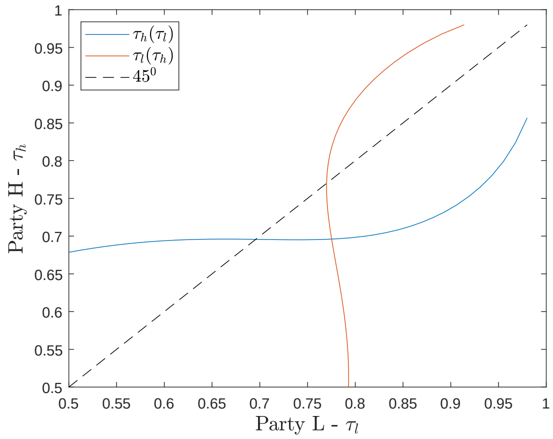
Citizen-office-seeking candidates. So far, candidates were simply interested in being elected. We now allow candidates to obtain utility from office but also to be interested in the policy outcome. Specifically, they suffer a loss that depends on the distance between their proposed tax rate and their bliss policy.[38] This has two effects on the equilibrium outcome. First, the proposed taxes by the two candidates are no longer equal. Second, there is activism in equilibrium.

Formally, the candidate from party H seeks to maximize the probability of getting elected, but would also like his bliss policy τ_h^* to be implemented in equilibrium. Given τ_l, this candidate solves the following program:

$$\max_{\tau_h} \mu p_h + (1-\mu) E_\tau \left[\frac{\left(\tau - \tau_h^*\right)^2}{2} \right], \tag{41}$$

where p_h is given by (33). $0 < \mu < 1$ is a preference weight for being elected relative to the loss incurred as the actual policy rate deviates from the candidate's bliss policy. The candidate from party L solves a similar program with bliss policy τ_l^*.

Figure 5: Citizen-Candidates Best Response



This figure represents the best response function of candidates at the platform choice stage, with $\mu = 1/2$, $\tau_l^* = 0$, $\tau_h^* = 1$. [Illustrative calibration: $\beta = 0.96$, $\alpha = \delta = 0.3$, all other parameters set to 1]

Figure 5 illustrates candidates' best responses and the equilibrium outcome: $\tau_l^p < \tau_h^p$ since candidates offer policy platforms that are now influenced by their individual preferences $\tau_l^* < \tau_h^*$. As explained in Section 4.3 and explicit in (32), different policy platforms are associated in equilibrium with conflictual activism along the equilibrium path.

5 Conclusions

Empirically, labor taxes are progressive but do not expropriate all the earnings of the rich. Instead, redistribution is limited. This paper argues that *group activism* is a factor that limits redistribution in a democracy. By activism, we mean influence campaigns that contribute to shape voters' preferences.

[38]Discussions with Annika Bacher led to the development of this case.

The analysis develops this mechanism in a dynamic setting in which the vote on taxes is taken after a human capital accumulation decision. Majority voting remains but the progress of political persuasion facilitates a redistribution of political power. Though there is no commitment, the outcome with activism is closer to the efficient allocation: redistribution is incomplete and incentives are retained for the accumulation of human capital.

But, there are conditions. First, the essence is in the timing not in the nature of group choice *per se*. Second activism must be a group decision. Coalitions of agents jointly decide on the level of activism and internalize the effects of their actions on the probabilities associated with the election outcome and thus on individual actions.

Throughout the analysis, we have refrained from consideration of alternative political institutions, such as ethical voting, that might coexist with activism. Understanding the relative importance of these mechanisms, perhaps in an empirical setting such as Coate and Conlin (2004), is a logical next step.

References

Acemoglu, D., M. Golosov, and A. Tsyvinski (2010): "Dynamic Mirrlees Taxation under Political Economy Constraints," The Review of Economic Studies, 77(3), 841–881.

Alesina, A., and G.-M. Angeletos (2005): "Fairness and Redistribution," American Economic Review, 95(4), 960–980.

Alesina, A., and E. L. Ferrara (2005): "Preferences for Redistribution in the Land of Opportunities," Journal of Public Economics, 89, 897–931.

Alesina, A., and N. Fuchs-Schündeln (2007): "Goodbye Lenin (or Not?): The Effect of Communism on People's Preferences," American Economic Review, 97(4), 1507–1528.

Alesina, A., and P. Giuliano (2011): "Chapter 4 - Preferences for Redistribution," vol. 1 of Handbook of Social Economics, pp. 93 – 131. North-Holland.

Alesina, A., and H. Rosenthal (1996): "A Theory of Divided Government," Econometrica, 64(6), 1311–41.

Alesina, A., S. Stantcheva, and E. Teso (2018): "Intergenerational Mobility and Preferences for Redistribution," American Economic Review, 108(2), 521–54.

Beck, P. A., R. J. Dalton, S. Greene, and R. Huckfeldt (2002): "The Social Calculus of Voting: Interpersonal, Media, and Organizational Influences on Presidential Choices," The American Political Science Review, 96(1), 57–73.

Becker, G. S. (1983): "A Theory of Competition Among Pressure Groups for Political Influence," The Quarterly Journal of Economics, 98(3), 371–400.

Bénabou, R. (2000): "Unequal Societies: Income Distribution and the Social Contract," American Economic Review, 90(1), 96–129.

——— (2002): "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?," Econometrica, 70(2), 481–517.

——— (2008): "Ideology," Journal of the European Economic Association, 6(2-3), 321–352.

Bénabou, R., and E. A. Ok (2001): "Social Mobility and the Demand for Redistribution: The Poum Hypothesis," The Quarterly Journal of Economics, 116(2), 447–487.

Bénabou, R., and J. Tirole (2006): "Belief in a Just World and Redistributive Politics," The Quarterly Journal of Economics, 121(2), 699–746.

Bierbrauer, F., A. Tsyvinski, and N. D. Werquin (2017): "Taxes and Turnout," Working Paper 24123, National Bureau of Economic Research.

Bierbrauer, F. J., and P. C. Boyer (2013): "Political Competition and Mirrleesian Income Taxation: A First Pass," Journal of Public Economics, 103, 1 – 14.

——— (2015): "Efficiency, Welfare, and Political Competition," The Quarterly Journal of Economics, 131(1), 461–518.

Casella, A. (2005): "Storable Votes," Games and Economic Behavior, 51(2), 391–419.

Casella, A., and A. Macé (2020): "Does Vote Trading Improve Welfare?," Working Paper 27743, National Bureau of Economic Research.

Caselli, F., and M. Morelli (2004): "Bad politicians," Journal of Public Economics, 88(3), 759 – 782.

Coate, S., and M. Conlin (2004): "A Group Rule-Utilitarian Approach to Voter Turnout: Theory and Evidence," American Economic Review, 94(5), 1476–1504.

DellaVigna, S., and M. Gentzkow (2010): "Persuasion: Empirical Evidence," Annual Review of Economics, 2(1), 643–669.

Farhi, E., C. Sleet, I. Werning, and S. Yeltekin (2012): "Nonlinear Capital Taxation without Commitment," Review of Economic Studies, 79(4), 1469–1493.

Feddersen, T., and A. Sandroni (2006): "A Theory of Participation in Elections," American Economic Review, 96(4), 1271–1282.

Fernández, R., and G. Levy (2008): "Diversity and Redistribution," Journal of Public Economics, 92(5-6), 925–943.

Fischer, S. (1980): "Dynamic Inconsistency, Cooperation and the Benevolent Dissembling Government," Journal of Economic Dynamics and Control, 2, 93–107.

Heathcote, J., K. Storesletten, and G. L. Violante (2017): "Optimal Tax Progressivity: An Analytical Framework," The Quarterly Journal of Economics, 132(4), 1693–1754.

Karabarbounis, L. (2011): "One Dollar, One Vote," The Economic Journal, 121(553), 621–651.

Levine, D. K., and A. Mattozzi (2020): "Voter Turnout with Peer Punishment," American Economic Review, 110(10), 3298–3314.

Lindbeck, A., and J. W. Weibull (1987): "Balanced-Budget Redistribution as the Outcome of Political Competition," Public Choice, 52(3), 273–297.

Maisel, L. S., J. M. Berry, G. C. Edwards, and K. L. Schlozman (2010): Who Sings in the Heavenly Chorus? The Shape of the Organized Interest System. Oxford University Press.

Mattozzi, A., and A. Merlo (2008): "Political careers or career politicians?," Journal of Public Economics, 92(3), 597 – 608.

Mattozzi, A., and E. Snowberg (2018): "The Right Type of Legislator: A Theory of Taxation and Representation," Journal of Public Economics, 159(C), 54–65.

MELTZER, A. H., AND S. F. RICHARD (1981): "A Rational Theory of the Size of Government," Journal of Political Economy, 89(5), 914–927.

OLSON, M. (1965): The Logic of Collective Action: Public Goods and the Theory of Groups., vol. 1. Harvard University Press, 1 edn.

PENNEC, C. L., AND V. PONS (2019): "Vote Choice Formation and Minimal Effects of TV Debates: Evidence from 61 Elections in 9 OECD Countries," Working Paper 26572, National Bureau of Economic Research.

PERSSON, T., AND G. TABELLINI (2002): Political Economics: Explaining Economic Policy, vol. 1. The MIT Press, 1 edn.

PIKETTY, T. (1995): "Social Mobility and Redistributive Politics," The Quarterly Journal of Economics, 110(3), 551–584.

ROEMER, J. E. (1998): "Why the poor do not expropriate the rich: an old argument in new garb," Journal of Public Economics, 70(3), 399 – 424.

SCHEUER, F., AND A. WOLITZKY (2016): "Capital Taxation under Political Constraints," American Economic Review, 106(8), 2304–28.

Mathematical Appendix

A Section 3 - Environment

Individual choice. Consider household optimization (1), substitute the constraints into the objective function:

$$\max_e \log(\theta - e) + \beta \int_\tau \Big\{(1-\tau)\big[\alpha \log(\theta) + \delta \log(e) - \frac{w^2}{2}\big] + \tau \log(\bar{\theta}')\Big\} dG(\tau), \tag{A.1}$$

where $G(\cdot)$ captures uncertainty over rate τ. The first order condition reads:

$$-\frac{1}{\theta - e} + \frac{\beta\delta(1-\bar{\tau})}{e} = 0, \tag{A.2}$$

where $\bar{\tau} = E(\tau)$. Reorganize and get (5).

Evolution of the income distribution. Start from (3), take the log and use (5):

$$\log(\theta') = \log(z) + (\alpha + \delta)\log(\theta) + \delta \log\big(\epsilon(\bar{\tau})\big). \tag{A.3}$$

The mean and variance of this expression yield (6) and (7).

Break-even income level. Given τ, income net of taxes and transfers satisfies (4), take the integral and then the log:

$$\log \int_\theta c' dF(\theta) = \tau \log(\bar{\theta}') + \log \int_\theta {\theta'}^{1-\tau} dF(\theta). \tag{A.4}$$

If $X \sim \log \mathcal{N}(\mu, \sigma^2)$, then $E(X^n) = e^{n\mu + n^2\sigma^2/2}$. Since $E(c') = E(\theta')$, it gives:

$$m' + \frac{{\sigma'}^2}{2} = \tau \log(\bar{\theta}') + (1-\tau)m' + \frac{(1-\tau)^2}{2}{\sigma'}^2. \tag{A.5}$$

Reorganize and get (8).

Value functions and bliss policies. Lifetime utility to an agent with initial income θ at $t = 1$ reads:

$$V_1(\theta, \tau) = v_1(\theta, e, \tau, \bar{\theta}') = \log(\theta - e) + \beta\big[(1-\tau)\big(\underbrace{\alpha \log(\theta) + \delta \log(e) - \frac{w^2}{2}}_{=E_z\left(\log(\theta')|\theta\right)}\big) + \tau \log(\bar{\theta}')\big], \tag{A.6}$$

where $e = \epsilon(\tau)\theta$ as in (5) and $\bar{\theta}'$ is given by (8). The first order condition w.r.t. τ:

$$\frac{dV_1(\cdot)}{d\tau} = \underbrace{\frac{\partial v_1(\cdot)}{\partial e}}_{=0}\frac{de}{d\tau} + \frac{\partial v_1(\cdot)}{\partial \tau} + \frac{\partial v_1(\cdot)}{\partial \log \bar{\theta}'}\frac{d \log \bar{\theta}'}{d\tau} = 0. \tag{A.7}$$

The first term is 0 from the envelope condition, the other terms are:

$$\frac{\partial v_1(\cdot)}{\partial \tau} = \log(\bar{\theta}') - \big(\alpha \log(\theta) + \delta \log(e) - \frac{w^2}{2}\big), \tag{A.8}$$

$$\frac{d \log(\bar{\theta}')}{d\tau} = \frac{dm'}{d\tau} - \frac{\sigma'^2}{2} = \delta \frac{\epsilon'(\cdot)}{\epsilon(\cdot)} - \frac{\sigma'^2}{2}. \tag{A.9}$$

Reorganize and get (10):

$$\beta(\alpha+\delta)\big(m - \log(\theta)\big) + \beta\big(\underbrace{(\alpha+\delta)^2\sigma^2 + w^2}_{=\sigma'^2}\big)(1-\tau) + \beta\delta \frac{\tau}{\epsilon(\tau)}\frac{d\epsilon(\tau)}{d\tau} = 0. \tag{A.10}$$

This expression implicitly defines $\tau^*(\theta)$. The first two terms form a decreasing linear function of τ, whose intercept is decreasing in θ. The third term, the elasticity of the saving rate to the redistribution rate τ, is decreasing in τ. Formally:

$$\frac{d}{d\tau}\Big(\frac{\tau}{\epsilon(\tau)}\frac{d\epsilon(\tau)}{d\tau}\Big) = -\Big[\frac{\tau}{1-\tau}\frac{\beta\delta}{[1+\beta\delta(1-\tau)]^2} + \frac{1}{(1-\tau)^2}\frac{1}{1+\beta\delta(1-\tau)}\Big] < 0, \tag{A.11}$$

which goes to $-\infty$ when τ goes to 1. Altogether there is a unique solution $\tau^*(\theta) < 1$ to (A.10), decreasing in θ.

The value function to an agent with initial income θ after $t = 1$ consumption and education choice reads:

$$\begin{aligned} V_2(\theta, \tau|\epsilon) &= v_2(\theta, e, \tau, \bar{\theta}'|\epsilon) \\ &= \beta\big[(1-\tau)\Big(\alpha\log(\theta) + \delta\log(e) - \frac{w^2}{2}\Big) + \tau\log(\bar{\theta}')\big] \end{aligned} \tag{A.12}$$

The difference with (A.6) is that the education choice is no longer sensitive to the redistribution rate τ and $e = \epsilon\theta$. The sensitivity of individual value functions to τ satisfies:

$$\begin{aligned} \frac{dV_2(\cdot)}{d\tau} &= \frac{\partial v_2(\cdot)}{\partial\tau} + \frac{\partial v_2(\cdot)}{\partial \log\bar{\theta}'}\frac{d\log\bar{\theta}'}{d\tau} \\ &= \beta(\alpha+\delta)\big(m - \log(\theta)\big) + \beta(\alpha+\delta)^2\big(\underbrace{(\alpha+\delta)^2\sigma^2 + w^2}_{=\sigma'^2}\big)(1-\tau) \end{aligned} \tag{A.13}$$

Bliss policy $\tau^d(\theta) \in [0, 1]$ is either an interior solution to $\frac{dV_2(\cdot)}{d\tau} = 0$ or $\tau^d(\theta) = 1$. Given the linear nature of (A.13), there is a unique bliss policy, ordered by initial income θ.

Proof Proposition 1. A benevolent planner with commitment solves (13). The first-order condition using (10) reads

$$\int_\theta \beta(\alpha+\delta)\big(m - \log(\theta)\big) + \beta\big(\underbrace{(\alpha+\delta)^2\sigma^2 + w^2}_{=\sigma'^2}\big)(1-\tau) + \beta\delta\frac{\tau}{\epsilon(\tau)}\frac{d\epsilon(\tau)}{d\tau} dF(\theta) = 0. \tag{A.14}$$

which gives (15).

Under a lack of commitment, it solves (14) given ϵ, which using (12) gives $\tau^d = 1$.

Sensitivity of optimal redistribution rate. To derive comparative statics for τ^*, first get

$$\frac{\tau}{\epsilon(\tau)}\frac{d\epsilon(\tau)}{d\tau} = -\frac{\tau}{1-\tau}\frac{1}{1+\beta\delta(1-\tau)}. \tag{A.15}$$

Then rewrite (15) as

$$\big((\alpha+\delta)^2\sigma^2 + w^2\big)(1-\tau)^2 - \frac{\delta\tau}{1+\beta\delta(1-\tau)} = 0 \tag{A.16}$$

The total derivative of this expression:

$$\begin{aligned}(\alpha+\delta)^2(1-\tau)^2 d\sigma^2 + (1-\tau)^2 dw^2 + \Big[2(\alpha+\delta)(1-\tau)^2 - \tau\frac{1+2\beta\delta(1-\tau)}{\big(1+\beta\delta(1-\tau)\big)^2}\Big]d\delta \\ = \Big[2{\sigma'}^2(1-\tau) + \delta\frac{1+\beta\delta}{\big(1+\beta\delta(1-\tau)\big)^2}\Big]d\tau\end{aligned} \tag{A.17}$$

Get immediately $\frac{d\tau^*}{d\sigma^2} > 0$ and $\frac{d\tau^*}{dw^2} > 0$ while the sign of $\frac{d\tau^*}{d\delta}$ is ambiguous.

B Section 4 - Probabilistic Voting and Activism

B.1 No Activism.

Comparison of platforms (19). Start from (11):

$$V_2(\theta,\tau|\epsilon) = \beta\Big((1-\tau)\big(\alpha\log(\theta) + \delta\log(\epsilon\theta) - \frac{w^2}{2}\big) + \tau\log(\bar{\theta}')\Big), \tag{B.1}$$

with $\log(\bar{\theta}')$ given by (8). Rearrange and get:

$$V_2(\theta,\tau|\epsilon) = \beta\Big[(1-\tau)(\alpha+\delta)\log(\theta) + \delta\log(\epsilon) - \frac{w^2}{2} + \tau(\alpha+\delta)m + \frac{{\sigma'}^2}{2}(2-\tau)\tau\Big]. \tag{B.2}$$

The difference of this expression with τ_h and τ_l:

$$\Delta V_2(\theta) = \beta(\tau_h-\tau_l)(\alpha+\delta)\big(m-\log(\theta)\big) + \beta\frac{{\sigma'}^2}{2}\big[(2-\tau_h)\tau_h - (2-\tau_l)\tau_l\big]. \tag{B.3}$$

Verify $(2-\tau_h)\tau_h - (2-\tau_l)\tau_l = (\tau_h-\tau_l)(2-\tau_h-\tau_l)$ and get (19). Finally, note

$$\int_\theta \Delta V_2(\theta)dF(\theta) = \beta(\tau_h-\tau_l)\frac{{\sigma'}^2}{2}(2-\tau_h-\tau_l). \tag{B.4}$$

Probability p_h of high rate of redistribution τ_h. Consider the following distributions: $\chi_\theta^j \sim U\Big(-\frac{1}{2\phi_\theta}+m, \frac{1}{2\phi_\theta}+m\Big)$ and $\psi \sim U\Big(-\frac{1}{2\Psi}+M, \frac{1}{2\Psi}+M\Big)$. Note $\phi = E(\phi_\theta)$. Given ψ and competing platforms (τ_l,τ_h), agent j with income θ votes for party H if and only if $\chi_\theta^j \leq \chi(\theta,\psi) = \Delta V_2(\theta) - \psi$. Hence, the share of agents with income

θ that vote for party H is:

$$\pi_{\theta,h}(\psi) = \int_{-\frac{1}{2\phi_\theta}+m}^{\chi(\theta,\psi)} \phi_\theta dj = \phi_\theta\Big(\chi(\theta,\psi) + \frac{1}{2\phi_\theta} - m\Big). \tag{B.5}$$

The share of votes across groups is then $\pi_h(\psi) = \int_\theta \pi_{\theta,h}(\psi)dF(\theta)$:

$$\pi_h(\psi) = \int_\theta \phi_\theta\Big(\chi_\theta + \frac{1}{2\phi_\theta} - m\Big)dF(\theta) = \int_\theta \phi_\theta \Delta V_2(\theta)dF(\theta) - \phi(\psi + m) + \frac{1}{2}. \tag{B.6}$$

The probability that party H wins the election is $p_h = P\big(\pi_h(\psi) \geq \lambda\big)$, where $\lambda \in [0,1]$ is a majority requirement for party H to win the election. The event $\pi_h(\psi) \geq \lambda$ is equivalent to the event

$$\psi \leq \bar{\psi} = \frac{1}{\phi}\int_\theta \phi_\theta \Delta V_2(\theta)dF(\theta) - m + \frac{1}{\phi}\big(\frac{1}{2} - \lambda\big). \tag{B.7}$$

Get then p_h as

$$p_h = P\big(\psi \leq \bar{\psi}\big) = \frac{1}{2} + \frac{\Psi}{\phi}\int_\theta \phi_\theta \Delta V_2(\theta)dF(\theta) - \Psi(m + M) + \frac{\Psi}{\phi}\Big(\frac{1}{2} - \lambda\Big). \tag{B.8}$$

Set $\phi_\theta = \phi$, $\lambda = \frac{1}{2}$, $m = M = 0$ and using (B.4) get (22).

Asymmetric political preferences. Persson and Tabellini (2002) study probabilistic voting allowing for a correlation between income and political preferences. To see how this mechanism works without activism, normalize average political preferences heterogeneity $\phi = E(\phi_\theta) = 1$ and assume that $\mathrm{cov}\big(\phi_\theta, \log(\theta)\big) = a$: if $a > 0$ then high income agents are more responsive to economic factors. The probability $p_h(\tau_l, \tau_h)$ is then

$$p_h(\tau_h, \tau_l) = \frac{1}{2} + \Psi\beta(\tau_h - \tau_l)\Big(-(\alpha + \delta)a + \frac{\sigma'^2}{2}(2 - \tau_h - \tau_l)\Big), \tag{B.9}$$

and a Nash equilibrium of economic platforms competition differs from full redistribution if and only if $a > 0$:

$$\tau_h = \tau_l = 1 - \frac{(\alpha + \delta)a}{\sigma'^2} < 1. \tag{B.10}$$

Overall, when high income agents are more sensitive to economic policy rather than political factors, they are attractive targets to candidates, which tilts equilibrium economic platform toward lower rate of redistribution.

B.2 No Activism, vote before the choice of education.

This section derives the equilibrium outcome of the game when the vote over competing policy platforms is taking place before individual education choices. The exposition follows Section 4.2.

Voting outcome. Given policy platforms (τ_l, τ_h), individuals trade off political and economic preferences and cast their vote. They evaluate policy platforms $\tau_l \leq \tau_h$ according to the value function $V_1(\theta, \tau)$ along with the realizations of idiosyncratic χ and aggregate ψ political preference shocks for candidate L. A type θ agent votes for party H if and only if:

$$V_1(\theta, \tau_h) > V_1(\theta, \tau_l) + \chi + \psi. \tag{B.11}$$

Let $\chi(\theta, \psi)$ be the *swing voter* for type θ: agents with income θ vote for party H if and only if $\chi \leq \chi(\theta, \psi)$, i.e.,

$$\chi(\theta, \psi) = V_1(\theta, \tau_h) - V_1(\theta, \tau_l) - \psi = \Delta V_1(\theta) - \psi, \tag{B.12}$$

where $\Delta V_1(\theta)$ is the economic gain (or loss) to agents with initial income θ of τ_h over τ_l. To derive this expression, start from (9) and note that at this stage m' is sensitive to τ. Rearrange and get:

$$V_1(\theta, \tau) = \log(\theta) + \log\big(1 - \epsilon(\tau)\big) + \beta\Big[(1-\tau)(\alpha+\delta)\log(\theta) + \delta \log\big(\epsilon(\tau)\big) - \frac{w^2}{2} + \tau(\alpha+\delta)m + \frac{\sigma'^2}{2}(2-\tau)\tau\Big] \tag{B.13}$$

Take the difference of when evaluated at τ_h and τ_l and using $(2-\tau_h)\tau_h - (2-\tau_l)\tau_l = (\tau_h - \tau_l)(2 - \tau_h - \tau_l)$, get:

$$\begin{aligned} \Delta V_1(\theta) =& \log\big(1-\epsilon(\tau_h)\big) - \log\big(1-\epsilon(\tau_l)\big) + \beta\delta\big[\log\big(\epsilon(\tau_h)\big) - \log\big(\epsilon(\tau_h)\big)\big] \\ &+ \beta(\tau_h - \tau_l)\Big[(\alpha+\delta)\big(m - \log(\theta)\big) + \frac{\sigma'^2}{2}(2-\tau_h-\tau_l)\Big]. \end{aligned} \tag{B.14}$$

The vote share for party H within group θ and across the population are given by (20), while the probability p_h that the candidate from party H wins the election is (21). Following the steps in Appendix B.1, one gets:

$$p_h = \frac{1}{2} + \Psi L(\boldsymbol{\tau}) + \Psi\beta(\tau_h - \tau_l)\frac{\sigma'^2}{2}(2 - \tau_h - \tau_l). \tag{B.15}$$

where $L(\boldsymbol{\tau}) = \log\big(1-\epsilon(\tau_h)\big) - \log\big(1-\epsilon(\tau_l)\big) + \beta\delta\big[\log\big(\epsilon(\tau_h)\big) - \log\big(\epsilon(\tau_h)\big)\big]$.

Choice of platforms. Given τ_l, a candidate from party H chooses τ_h to maximize the probability (B.15) of winning the election. The first order condition is:

$$\frac{dp_h(\cdot)}{d\tau_h} = \Psi\beta\Big[\sigma'^2(1-\tau_h) + \delta\frac{\tau_h}{\epsilon(\tau_h)}\frac{d\epsilon(\tau_h)}{d\tau_h}\Big] = 0, \tag{B.16}$$

where we use

$$\frac{dL(\boldsymbol{\tau})}{d\tau_h} = \frac{d\epsilon(\tau_h)}{d\tau_)}\Big[\frac{\beta\delta}{\epsilon(\tau_h)} - \frac{1}{1-\epsilon(\tau_h)}\Big] = \beta\delta\frac{\tau_h}{\epsilon(\tau_h)}\frac{d\epsilon(\tau_h)}{d\tau_h}. \tag{B.17}$$

Condition (B.16) coincides with (15): $\tau_h = \tau^*$ is a dominant strategy. The candidate from party L maximizes $p_l(\tau_l, \tau_h) = 1 - p_h(\tau_l, \tau_h)$, and again $\tau_l = \tau^*$ is a dominant strategy. The outcome of the game is the efficient tax rate $\tau^p = \tau^*$.

B.3 Electoral Competition under Activism

Activism choice. To get (30), rewrite (28) as:

$$V_1(\theta, \boldsymbol{\tau}) = v_1\big(\theta, \boldsymbol{\tau}, e, p_h, \log(\bar{\theta}'_h), \log(\bar{\theta}'_l)\big), \tag{B.18}$$

where $e = \epsilon(\bar{\tau})\theta$. Further, $\frac{dV_1(\cdot)}{dA^i_\theta} = \frac{dV_1(\cdot)}{dp_h}\frac{dp_h}{dA^i_\theta}$. Consider the first term:

$$\frac{dV_1(\cdot)}{dp_h} = \underbrace{\frac{\partial v_1(\cdot)}{\partial e}}_{=0}\frac{de}{dp_h} + \frac{\partial v_1(\cdot)}{\partial p_h} + \sum_{i=l,h}\frac{\partial v_1(\cdot)}{\partial \log(\bar{\theta}'_i)}\frac{d\log(\bar{\theta}'_i)}{dp_h}. \tag{B.19}$$

Term by term:

$$\frac{\partial v_1(\cdot)}{\partial p_h} = -\beta(\tau_h - \tau_l)E_z\big(\log(\theta')|\theta\big) + \beta\big(\tau_h\log(\bar{\theta}'_h) - \tau_l\log(\bar{\theta}'_l)\big), \tag{B.20}$$

and

$$\tau_h\log(\bar{\theta}'_h) - \tau_l\log(\bar{\theta}'_l) = (\tau_h - \tau_l)\Big[m' + \frac{{\sigma'}^2}{2}(2 - \tau_h - \tau_l)\Big]. \tag{B.21}$$

Using (6) and $E_z\big(\log(\theta')|\theta\big) = (\alpha+\delta)\log(\theta) + \delta\log\big(\epsilon(\bar{\tau})\big) - \frac{w^2}{2}$, rearrange and get:

$$\frac{\partial v_1(\cdot)}{\partial p_h} = \beta(\tau_h - \tau_l)\Big[(\alpha+\delta)\big(m - \log(\theta)\big) + \frac{{\sigma'}^2}{2}(2 - \tau_h - \tau_l)\Big]. \tag{B.22}$$

Then, since $\frac{\partial v_1(\cdot)}{\partial \log(\bar{\theta}'_i)} = \beta p_i \tau_i$ and $\frac{d\log(\bar{\theta}'_i)}{dp_h} = \delta\frac{d\bar{\tau}}{dp_h}\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})}$ for $i \in (l, h)$,

$$\sum_{i=l,h}\frac{\partial v_1(\cdot)}{\partial \log(\bar{\theta}'_i)}\frac{d\log(\bar{\theta}'_i)}{dp_h} = \beta(\tau_h - \tau_l)\delta\underbrace{\big(p_h\tau_h + (1-p_h)\tau_l\big)}_{=\bar{\tau}}\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})}. \tag{B.23}$$

Overall, with $\frac{dp_h}{dA^i_\theta} = \pm\Psi\gamma$, with $\pm = +$ for $x = h$ and $\pm = -$ for $x = l$, one gets (30):

$$\frac{dV_1(\cdot)}{dA^i_\theta} = \pm\Psi\gamma\beta(\tau_h - \tau_l)\Big[(\alpha+\delta)\big(m - \log(\theta)\big) + \frac{{\sigma'}^2}{2}(2 - \tau_h - \tau_l) + \delta\bar{\tau}\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})}\Big]. \tag{B.24}$$

Lemma 1. Given policy platforms (τ_l, τ_h), an equilibrium of the activism subgame is a set of income group contributions $\{A^l_\theta, A^h_\theta\}$, aggregate contributions (A^l, A^h) defined as $A^i = \int_\theta A^i_\theta d\theta$, probability of high taxes p_h and expected tax rate $\bar{\tau} = E(\tau)$.

To establish existence, start from (29) characterizing optimal contribution of income groups. Sum this expression over income group and get the aggregate effects of activism on political preferences (32).

Using this expression with (26), get (33) which uniquely defines the probability p_h, since the right-hand side of this expression is decreasing in $\bar{\tau}$, hence in p_h. From this probability, one can recover aggregate and individual activism contributions. The probability p_h is unique, because the sensitivity of the education rate to the expected redistribution rate $\bar{\tau}\frac{\epsilon'(\cdot)}{\epsilon(\cdot)} = -\frac{\bar{\tau}}{1-\bar{\tau}}\frac{1}{1+\beta\delta(1-\bar{\tau})} < 0$ depends negatively on $p_h(\tau_l, \tau_h)$, since $\bar{\tau} = p_h\tau_h + (1-p_h)\tau_l$.

The argument for uniqueness goes as follow. Given (τ_h, τ_l), the outcome of the contribution game yield a unique p_h (26). Is there another pair (A^h, A^l) and underlying group contributions that yield the same outcome? No, because the marginal return to contribution is a function of $A^h - A^l$, but the marginal cost depends on individual contribution only, see (29).

Endogenous choice of platforms. To derive (35), denote by $H(\tau_h, \tau_l, \bar{\tau})$ the right hand side of (33) and let $\mathcal{E}(\bar{\tau}) = \bar{\tau}\frac{\epsilon'(\bar{\tau})}{\epsilon(\bar{\tau})}$ be the elasticity of education rate to the expected tax rate. Totally differentiating (33) w.r.t. p_h and τ_h:

$$\Big[1 - \frac{\partial H(\cdot)}{\partial \bar{\tau}}\frac{d\bar{\tau}}{dp_h}\Big]dp_h = \Big[\frac{\partial H(\cdot)}{\partial \tau_h} + \frac{\partial H(\cdot)}{\partial \bar{\tau}}\frac{d\bar{\tau}}{d\tau_h}\Big]d\tau_h. \tag{B.25}$$

Term by term:

$$\frac{\partial H(\cdot)}{\partial \tau_h} = \Psi\beta\big[(1+\Psi\gamma^2)\sigma'^2(1-\tau_h) + \Psi\gamma^2\delta\mathcal{E}(\bar{\tau})\big], \tag{B.26}$$

$$\frac{\partial H(\cdot)}{\partial \bar{\tau}} = \Psi\beta(\tau_h - \tau_l)\Psi\gamma^2\delta\frac{d\mathcal{E}(\bar{\tau})}{d\bar{\tau}}, \tag{B.27}$$

$$\frac{d\mathcal{E}(\bar{\tau})}{d\bar{\tau}} = -\frac{1+\beta\delta-\beta\delta\tau^2}{(1-\tau)^2\big[1+\beta\delta(1-\tau)\big]^2}. \tag{B.28}$$

(B.29)

Rearranging terms, get (35).

Proposition 3. If $\gamma = 0$ or $\delta = 0$, then the unique solution to (36) is $\tau = 1$. Otherwise, using (A.15), rewrite (36) as:

$$\Big(1 + \frac{1}{\Psi\gamma^2}\Big)\sigma'^2(1-\tau) = \frac{\tau}{1-\tau}\frac{\delta}{1+\beta\delta(1-\tau)}. \tag{B.30}$$

The left hand side is decreasing in τ, while the right hand side is increasing in τ, with limit when $\tau = 1$ is ∞, which gives that the unique solution satisfies $0 < \tau^p < 1$. An increase in γ decreases the left hand side, which yield $\frac{d\tau^p}{d\gamma} < 0$. In the limit $\gamma = \infty$, (36) coincides with (15), hence $\tau^p = \tau^*$.

B.4 Discrete partition of income groups and activism.

This appendix recasts the activism game presented in Section 4.3 to account for a discrete numbers of influence groups.

After the announcement of policy platforms, the population is partitioned in $N \geq 2$ groups ordered by income level θ: $\theta_l = \theta_1 < \theta_2 < ... < \theta_h = \theta_{N+1}$. Each group $j \in \{1, 2, ...N\}$ is composed of agents $[\theta_j, \theta_{j+1}]$ and decides on activism intensity $A^h_j \geq 0$ and $A^l_j \geq 0$ to promote candidates and their policy platform. Note $A^i = \sum_{j=1}^N A^i_j$ the aggregated influence of activism for each candidate $i \in \{h, l\}$. The probability p_h that party H wins the election is given by (26).

Given competing platforms $\boldsymbol{\tau} = (\tau_l, \tau_h)$ and all other group activism decisions $\{A^l_{-j}, A^h_{-j}\}$, income group j decides on total group contributions $\big(A^l_j, A^h_j\big)$:

$$\max_{A^l_j, A^h_j \geq 0} \int_{\theta_j}^{\theta_{j+1}} f(\theta) V_1(\theta, \boldsymbol{\tau}) d\theta - \frac{1}{2}\Big((A^l_j)^2 + (A^h_j)^2\Big). \tag{B.31}$$

The first order condition for $A^i_j \geq 0$, $i \in (l, h)$, is

$$\int_{\theta_j}^{\theta_{j+1}} f(\theta) \frac{dV_1(\theta)}{dA^i_\theta} d\theta = A^i_j. \tag{B.32}$$

Assume that the partition is such that every group is composed of agents that all favors the same candidate, i.e., there is a $n \in \{2, ..., N\}$ s.t. $\theta_n = \hat{\theta}$ defined in (31). Aggregating (B.32) across groups j, one gets (32), and the rest of the analysis is unchanged.

B.5 Citizen-office-seeking candidate

The program of the candidate from party H

$$\max_{\tau_h} \mu p_h + (1-\mu)\Big[p_h \frac{(\tau_h - \tau_h^*)^2}{2} + (1-p_h)\frac{(\tau_l - \tau_h^*)^2}{2}\Big] \tag{B.33}$$

$$\mu \frac{dp_h}{d\tau_h} + (1-\mu)\Big[\frac{dp_h}{d\tau_h}\frac{(\tau_h - \tau_h^*)^2}{2} + p_h(\tau_h - \tau_h^*) - \frac{dp_h}{d\tau_h}\frac{(\tau_l - \tau_h^*)^2}{2}\Big] = 0 \tag{B.34}$$

Similarly, the first order condition characterizing $\tau_l(\tau_h)$:

$$-\mu \frac{dp_h}{d\tau_l} + (1-\mu)\Big[\frac{dp_h}{d\tau_l}\frac{(\tau_h - \tau_l^*)^2}{2} + (1-p_h)(\tau_l - \tau_l^*) - \frac{dp_h}{d\tau_l}\frac{(\tau_l - \tau_l^*)^2}{2}\Big] = 0 \tag{B.35}$$

To derive these expressions, denote $p_h = G(\tau_l, \tau_h, \bar{\tau})$ and get:

$$\Big[1 - \frac{\partial G(\cdot)}{\partial \bar{\tau}}\frac{d\bar{\tau}}{dp_h}\Big] dp_h = \Big[\frac{\partial G(\cdot)}{\partial \tau_i} + \frac{\partial G(\cdot)}{\partial \bar{\tau}}\frac{d\bar{\tau}}{d\tau_i}\Big] d\tau_i \tag{B.36}$$

Then derive

$$\frac{\partial G(\cdot)}{\partial \bar{\tau}} = \Psi\beta(\tau_h - \tau_l)\Psi\gamma^2\delta\frac{d\mathcal{E}(\bar{\tau})}{d\bar{\tau}} \qquad \frac{d\bar{\tau}}{dp_h} = \tau_h - \tau_l \tag{B.37}$$

$$\frac{d\bar{\tau}}{d\tau_h} = p_h \qquad \frac{d\bar{\tau}}{d\tau_l} = (1-p_h) \tag{B.38}$$

and

$$\frac{\partial G(\cdot)}{\partial \tau_h} = \Psi\beta\big[(1+\Psi\gamma^2){\sigma'}^2(1-\tau_h) + \Psi\gamma^2\delta\mathcal{E}(\bar{\tau})\big] \tag{B.39}$$

$$\frac{\partial G(\cdot)}{\partial \tau_l} = -\Psi\beta\big[(1+\Psi\gamma^2){\sigma'}^2(1-\tau_l) + \Psi\gamma^2\delta\mathcal{E}(\bar{\tau})\big] \tag{B.40}$$

Finally

$$\mathcal{E}(\tau) = -\frac{\tau}{1-\tau}\frac{1}{1+\beta\delta(1-\tau)} \qquad\qquad \frac{d\mathcal{E}(\tau)}{d\tau} = -\frac{1+\beta\delta(1-\tau^2)}{(1-\tau)^2\big[1+\beta\delta(1-\tau)\big]^2} \tag{B.41}$$

C A Simple Framework: Timing and Group Choice

This section proposes simple elements to complement the exposition presented in the analysis. Section C.1 considers the implication of alternative timings on the effect of group choice on the selection of tax and transfers. Section C.2 highlights the differences between group *activism* and the *ethical voter* model, another popular group choice protocol.

C.1 The influence of group choice under different timings

Consider a simple consumption - labor choice set-up in a homogeneous agent economy, where group choice can influence individual preferences over tax rates. We contrast the following timings: taxes are chosen before or after the labor supply decision. This difference highlights the precise influence of group choice a on voters' tax preferences, and eventually the role of group choice in shaping equilibrium outcomes.

This simple abstract set up accommodates several interpretations of group choice: it can influence the outcome of the vote through an extensive margin - getting more people to vote, as in *ethical voter models* - or through influencing their preferences for low tax rates, as in *activism*.[39]

Timing 1: Platform - Group choice - Vote - Work

Study this game under backward induction.

Work. The consumption - labor choice is:

$$V(\theta,\tau) = \max_n u(\theta n(1-\tau)+T,n) \tag{C.1}$$

Note $n(\theta,\tau,T)$ the labor supply policy function. The government budget constraint is $T=\theta n\tau$, so that

$$\frac{dV(\theta,\tau)}{d\tau} = u_1\big(\theta n(1-\tau)+T,n\big)\Big(\frac{dT}{d\tau}-\theta n\Big) = u_1\big(\theta n(1-\tau)+T,n\big)\theta\tau\frac{dn(\cdot)}{d\tau}. \tag{C.2}$$

Note that labor supply depend on taxes, but not on the level of group choice a.

Vote. In this stylized game, the outcome of the vote is captured by $p(a)$, the probability of high tax outcome given group choice $a\in\mathbb{R}$. Assume $p'(a)<0$ for $a\neq 0$ and $p'(0)=0$..

[39]As mentioned in Section 2, analyses with *ethical voters* rely on the first timing while our analysis of *activism* considers the second timing.

Group choice. Under this timing, group choice a is taken prior to the vote:

$$\max_{a\in\mathbb{R}} \; p(a)V(\theta,\tau_h) + (1-p(a))V(\theta,\tau_l) - k(a) \tag{C.3}$$

where $k(a)$ is a symmetric cost function: $k(0) = 0$ and $k(a) = k(-a)$. The first order condition of this program yields:

$$p'(a)[V(\theta,\tau_h) - V(\theta,\tau_l)] = k'(a). \tag{C.4}$$

Clearly if $\tau_h > \tau_l$, then $a > 0$, i.e., the group choice is pushing for lower taxes. Under this timing, group choice a influences individual exposure to redistribution and distortions.

Platform. Under this timing, with or without group choice, $\tau^* = 0$ is the outcome of electoral competition: homogenous agents have no gain from redistribution, and distortionary taxation would only impose welfare losses.

Timing 2: Platform - Group Choice - Work - Vote

This alternative timing reflects the tension between individual long term investment decision and repeated elections that motivates our study.

Vote. Same as before, the outcome of the vote is determined by $p(a)$, the probability of high taxes, where $a \in \mathbb{R}$ is the group choice.

Work. Labor supply is taken with uncertainty over taxes $\boldsymbol{\tau} = (\tau_l, \tau_h)$

$$W(\theta,\boldsymbol{\tau},a) = \max_{n} \; p(a)u\big(\theta n(1-\tau_h) + T_h, n\big) + \big(1-p(a)\big)u\big(\theta n(1-\tau_l) + T_l, n\big) \tag{C.5}$$

The solution is given by $n = \phi(\theta,\boldsymbol{\tau},a)$. Note that the transfer is indexed in the same way as the *ex post* tax due to the government budget constraint: $T_j = \tau_j\theta\phi(\theta,\boldsymbol{\tau},a)$.

Group choice a influences the probability of high taxes and, through this, labor supply and output, as is explicit from the presence of a in the labor supply decision rule. This is the channel that is missing in the alternative timing and that allows group choice to influence the equilibrium outcome.

Group Choice. The choice of a is determined by

$$\max_{a} W(\theta,\boldsymbol{\tau},a) - k(a) \tag{C.6}$$

yielding:

$$\frac{dW(\theta,\boldsymbol{\tau},a)}{da} = k'(a). \tag{C.7}$$

Using (C.5), it becomes

$$\begin{aligned} & p'(a)\big[u\big(c_h,\phi(\theta,\boldsymbol{\tau},a)\big)-u\big(c_l,\phi(\theta,\boldsymbol{\tau},a)\big)\big] \\ & +\big[p(a)u\big(c_h,\phi(\theta,\boldsymbol{\tau},a)\big)\tau_h+\big(1-p(a)\big)u\big(c_l,\phi(\theta,\boldsymbol{\tau},a)\big)\tau_l\big]\theta\frac{d\phi(\cdot)}{da}=k'(a). \end{aligned} \tag{C.8}$$

where $u\big(c_j,\phi(\cdot)\big) = u\big(\theta\phi(\theta,\boldsymbol{\tau},a)(1-\tau_j)+T_j,\phi(\cdot)\big)$ is the realized level of utility in state $j = h,l$ with labor $n=\phi(\theta,\boldsymbol{\tau},a)$ determined before the realization of uncertainty on the tax level.

As under the alternative timing, the direct effect of group choice a on employment disappears by the envelope condition. The first term in this expression is similar to (C.3): it reflects the exposure to redistribution. But given the timing and the assumption of homogeneous agents, redistribution when implemented is not distortionary, and (C.8) simplifies to:

$$\big[p(a)u\big(c_h,\phi(\theta,\boldsymbol{\tau},a)\big)\tau_h+\big(1-p(a)\big)u\big(c_l,\phi(\theta,\boldsymbol{\tau},a)\big)\tau_l\big]\theta\frac{d\phi(\cdot)}{dp}\frac{dp(\cdot)}{da}=k'(a). \tag{C.9}$$

This second set of terms come from the effect of variations in a on the tax base, which influences the level of tax and transfers in state j. Under this timing, group choice contributes to preserve the tax base against the threat of excessive redistribution.

Platform. This example illustrates how group choice influences the equilibrium outcome and is sufficient to eliminate full redistribution. At the time of the vote, absent the effect of group choice on the probability of high taxes, i.e., $p'(a) = 0$ for all a, homogeneous agents are indifferent between tax platforms: any $\tau \in [0,1]$ and associated transfers can be an equilibrium of this game. When group choice is effective to influence the probability of high taxes, then (C.9) makes clear that full redistribution cannot be an outcome of electoral competition. Suppose that $(\tau_l,\tau_h)=(1,1)$ are equilibrium platforms. Then the group choice $a<0$ will make sure that deviation $\tau_l = 1-\epsilon$ wins the election, because of its beneficial effect on the tax base $\frac{d\phi(\cdot)}{da}=\frac{d\phi(\cdot)}{dp}\frac{dp(\cdot)}{da}$.

C.2 Ethical voter vs. activism: Timing 2

We consider a generic probabilistic voting model as in Section 4, under a common timing: Platform - Group Choice - Work - Vote. We expose explicitly how *activism* differs from *ethical voter* protocol. The latter operates through the composition of the voting population, while the former operates via its influence on the preferences of the voting population.

An agent with income θ and idiosyncratic political preference χ votes for high taxes τ_h given aggregate shock ψ and activism $\Delta A = A_l - A_h$ if and only if:

$$\chi \le \chi(\theta,\psi,\Delta A) = \Delta V_2(\theta)-\psi-\gamma\Delta A.$$

The share of agents with income θ that vote for high taxes:

$$\pi_{\theta,h}(\psi,\Delta A) = G_\chi\big(\Delta V_2(\theta)-\psi-\gamma\Delta A\big),$$

where $G\chi(\cdot)$ is the cumulative distribution function of idiosyncratic shocks χ. The share of the overall population that votes for high taxes is

$$\pi_h\big(\psi, \Delta A, x(\theta)\big) = \int_\theta \pi_{\theta,h}(\cdot)x(\theta)dF(\theta) = \int_\theta G_\chi\big(\Delta V_2(\theta) - \psi - \gamma\Delta A\big)x(\theta)dF(\theta),$$

where $x(\theta)$ is the share of agents with income θ that votes. The probability that the high tax platform wins the election is then:

$$p_h = P\big(\pi_h(\cdot) \geq 1/2\big)$$

This simple formulation highlights how the political protocols operate on the election outcome given competing tax platforms (τ_l, τ_h):

- Activism operates at the intensive margin, since it affects the preferences of the overall population: an increase in ΔA decreases the share of voters that vote for high taxes across the income distribution θ.

- The ethical voter protocol operates at the extensive margin, because it influences the composition of the voting population: if $x(\theta)$ is increasing in θ, then the probability of high taxes unambiguously go down, because relatively richer agent participate more into a vote about a pure redistributive conflict.